# PATENT ABSTRACTS OF JAPAN

(11)Publication number : **10-210409**
(43)Date of publication of application : **07.08.1998**

(51)Int.Cl. **H04N 5/92**
**G10L 9/18**
**H04N 5/907**

(21)Application number : **09-005844**
(22)Date of filing : **16.01.1997**

(71)Applicant : **SONY CORP**
(72)Inventor : **KOBAYASHI TOSHIHARU**

## (54) IMAGE AND VOICE RECORDERAND IMAGE AND VOICE RECORDER AND REPRODUCER USING SEMICONDUCTOR MEMORY

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a recorder and a recorder/reproducer which can secure synchronization between dynamic images and voices by recording the encoded and compressed image and voice data on an IC memory cardtogether with the time information that secures the synchronization between those image and voice data.
SOLUTION: Voice data are encoded at a voice compression part of a voice data compression/expansion circuit 4 and recorded in a memory 14consisting of a DRAMetc. Then the compressed voice data are read out the memory 14 via a card I/F circuit 15 and recorded on an IC memory card 16. The luminance and color difference signals produced from the dynamic image signals are encoded at an image compression part of an image compression/expansion circuit 13 and recorded in the memory 14. The compressed image data are read out of the memory 14 via the circuit 15 and recorded on the card 16. Furthermorethe time management information is recorded on the card 16to secure synchronization between the compressed image and the voice data.

## CLAIMS

[Claim(s)]
[Claim 1]A picture and a voice recording device characterized by comprising the following using semiconductor memory which carries out record reproduction of image data and voice data which were inputted.
An image encoding means which performs compression encoding to described image dataand generates variable-length coded image data.
A voice coding means to perform compression encoding to the above-mentioned

voice dataand to generate variable-length coded voice data.
A recording device which records the above-mentioned coded image data and the above-mentioned coded voice data on removable semiconductor memory.

[Claim 2]A picture and a voice recording device using semiconductor memory also recording simultaneously a hour entry for synchronizing the above-mentioned coded image data and the above-mentioned coded voice data with the above-mentioned semiconductor memory in claim 1.
[Claim 3]A picture and a voice recording device using semiconductor memory recording the above-mentioned coded image data and the above-mentioned coded voice data in DOS file form to the above-mentioned semiconductor memory in claim 1.
[Claim 4]A picture and voice recording playback equipment characterized by comprising the following using semiconductor memory which carries out record reproduction of image data and voice data which were inputted.
An image encoding means which performs compression encoding to described image dataand generates variable-length coded image data.
A voice coding means to perform compression encoding to the above-mentioned voice dataand to generate variable-length coded voice data.
A recording device which records the above-mentioned coded image data and the above-mentioned coded voice data on removable semiconductor memory.
An image decoding means which reads the above-mentioned coded image data from the above-mentioned semiconductor memorydecodes coded image data by which reading appearance was carried out [ above-mentioned ]and generates reproduced image dataA speech decoding means which reads the above-mentioned coded voice data from the above-mentioned semiconductor memorydecodes coded voice data by which reading appearance was carried out [ above-mentioned ]and generates playback voice data.

[Claim 5]A picture and voice recording playback equipment using semiconductor memory characterized by having displayed the above-mentioned reproduced image data on a display monitorand the above-mentioned playback voice data's taking the above-mentioned reproduced image data and a synchronizationand making it output them from a loudspeaker in claim 4.
[Claim 6]A picture and a voice recording device using semiconductor memory also recording simultaneously a hour entry for synchronizing the above-mentioned coded image data and the above-mentioned coded voice data with the above-mentioned semiconductor memory in claim 4.
[Claim 7]A picture and a voice recording device using semiconductor memory recording the above-mentioned coded image data and the above-mentioned coded voice data in DOS file form to the above-mentioned semiconductor memory in claim 4.

## DETAILED DESCRIPTION

[Detailed Description of the Invention]
[0001]
[Field of the Invention]This invention relates to the picturevoice recording deviceand recording and reproducing device using the semiconductor memory which can make easy record and reproduction of the data in the case of compressing an unfixed length method into video and a sound.
[0002]
[Description of the Prior Art]When recording the sound in sync with video and its video conventionallya means to record dynamic image data and voice data by time sharing (interleave) was usual. This is a general method seen with MPEG (Moving Picture Experts Group) 1 and 2 and digital VCR (video cassette recorder).
[0003]In MPEGas shownfor example in drawing 8 Awhen multiplexing dynamic image data and voice datamultiplexing by a packet is performed. For examplewhen carrying out multiplex [ of the dynamic image data ]dynamic image data and voice data are divided into the stream of the suitable length called a packetrespectivelyadditional informationsuch as a headeris attachedthe packet of dynamic image data and voice data is switched suitablyand time division transmission is made to be carried out. Actuallyas shown in drawing 8 Bdynamic image data and voice data are treated by the constitutional unit called the pack which bundled two or more packets. In that casethe additional information for referring to the time base for synchronous reproductionetc. are added to the portion of a pack header.
[0004]In digital VCRas shown in drawing 9 as an examplethe image data recorded as a slanting track on magnetic tapeCompression by DCT (discrete cosine transformation) is carried outand two kinds of error correcting codesoutside numerals and an inner codeare used togethershuffling for one scanning line before generating outside numeralsand shuffling in 1 sector before generating an inner code after adding outside numerals are performedand it is recorded on a recording medium. 204-block image data is recorded on one track. This block comprises a 2 words synchronous codea 2 words ID codean 85 words image code8 words inner code parityan 85 words image codeand 8-word inner code parity. The voice data can record the numerals up to 20 bits sampled at 48 kHzserves as the almost same composition as dynamic image dataand as shown in a figureit is recorded on the both ends of dynamic image data by the recording medium via a gap.
[0005]
[Problem(s) to be Solved by the Invention]Thuswhen video and a sound were recorded on a disk or a tapein the disc medium represented by an optical discHDDetc.it could not say that seek time was quick enoughand there was a problem to which seek time becomes very long in tape media to the extent that even the concept of seeking cannot be found. in order for there to be such restrictions and to make processing of a regenerative signal easy -- real time -- and each data is sequentially multiplexed and recorded by fixed length inevitably.

[0006]On the other handsince seek time is dramatically as high-speed as there is no concept of seeking itselfthe IC memory card does not need to manage dynamic image data and voice data as fixed length data. Since a scene thru/or a soundless state with little data volume also exists in order to make data volume into the minimumit is rational to adopt an unfixed length method as the compressing method. When an unfixed length method is adoptedif it is only the continuous reproduction from a headtrouble will not be produced by the method of managing video and a sound as another fileeitherbut if it performs that it is equivalent to the reproduction from the middleand rewinding to the middlea certain hour entry is required.
[0007]Thereforethe purpose of this invention is to provide the picturevoice recording deviceand recording and reproducing device using the semiconductor memory which can take the synchronization with video and a sound as a recording medium using semiconductor memorysuch as an IC memory card.
[0008]
[Means for Solving the Problem]In a picture and a voice recording device using semiconductor memory to which the invention according to claim 1 carries out record reproduction of image data and voice data which were inputtedAn image encoding means which performs compression encoding to image data and generates variable-length coded image dataIt is a picture and a voice recording device using semiconductor memory consisting of a voice coding means to perform compression encoding to voice data and to generate variable-length coded voice dataand a recording device which records coded image data and coded voice data on removable semiconductor memory.
[0009]In a picture and voice recording playback equipment using semiconductor memory to which the invention according to claim 4 carries out record reproduction of image data and voice data which were inputtedAn image encoding means which performs compression encoding to image data and generates variable-length coded image dataA voice coding means to perform compression encoding to voice data and to generate variable-length coded voice dataA recording device which records coded image data and coded voice data on removable semiconductor memoryAn image decoding means which reads coded image data from semiconductor memorydecodes read coded image dataand generates reproduced image dataIt is an image and voice recording playback equipment using semiconductor memory consisting of a speech decoding means which reads coded voice data from semiconductor memorydecodes read coded voice dataand generates playback voice data.
[0010]It is codedfor example by MPEG and dynamic image data is memorized by IC memory card by GOP unitsSynchronizing with the GOP unitsit is codedfor example by ATRACvoice data is memorized by IC memory cardand hour entry management for taking a synchronization with compressed image data and compression audio data further is memorized. By thiswhen reproducing arbitrary positionscompression audio data in sync with compressed image data and its compressed image data can be arbitrarily read from hour entry managementand it can reproduce.

[0011]
[Embodiment of the Invention]Hereafterone example of this invention is described in detail with reference to drawings. The composition of one example of this invention is shown in drawing 1. A sound is inputted from the microphone shown by 1. The inputted sound is changed into an electrical signal and supplied to the amplifier 2 as an audio signal by the sound / electrical signal sensing element contained in the microphone 1. In the amplifier 2an audio signal is amplified and it is restricted to a required zone. Thenan audio signal is digitized by a zone more than twice the sampling frequency of being required in the A/D conversion circuit 3. It is coded by the voice compression section of voice data compression / expansion circuit 4and this voice data is recorded on the memory 14 which consists of DRAMs etc.
[0012]The compression audio data recorded on the memory 14 is read via the card I/F circuit 15and is recorded with IC memory card 16. This IC memory card 16 is one of the storages which can be provided with the function of a common memory cardand can be written with a personal computer. As for IC memory card 16record/reproduction is made in the format of DOS form as an example. The connector which equips with IC memory card 16 between the card I/F circuit 15 and IC memory card 16 is allotted actuallyand attachment and detachment of an IC memory card are enabled by the connector.
[0013]The compression audio data recorded on IC memory card 16 is read via the card I/F circuit 15. The read compression audio data is written in the record section for the sounds of the memory 14. The written-in compression audio data is elongated by real time in the voice expanding part of voice data compression / expansion circuit 4. The elongated voice data is supplied to D/A converter 5and is analog-ized. The analog-ized audio signal is amplified with the amplifier 6and is outputted from the loudspeaker 7 as an audio signal.
[0014]The photographic subject used as video is inputted from a lens and the converging section 8. A photographic subject is inputted into CCD image sensor 9 from a lens and the converging section 8. In CCD image sensor 9a photographic subject is accumulated as an electric charge and changed into an electrical signal. An electrical signal is supplied to CDS circuit 10 as a dynamic image signal. CDS circuit 10 is a correlation double sampling circuitand aims at reduction of a noise. moreover -- CPU -- 22 -- a lens -- a converging section -- eight -- a mechanical system -- a driver -- 18 -- a CCD image sensor -- nine -- a CDS circuit -- ten -- and -- an A/D converter -- 11 -- an electric system -- a timing signal generation part -- (-- TG --) -- 19 -- controlling .
[0015]The output signal of CDS circuit 10 is supplied to A/D converter 11and is digitized with A/D converter 11 by a sampling frequency which serves as an integral multiple of the optimal sampling frequencyfor examplethe subcarrier frequency of an NTSC signal. The output signal of digitized CDS circuit 10 is supplied to the digital signal processing circuit 12. In the digital signal processing circuit 12the usual camera signal processingsuch as a gamma correction and color separationis made to the supplied signaland the luminance signal Ythe color-

difference signal Cband Cr are created. It is coded by the image compression part of image data compression / expansion circuit 13and these signals are recorded on the memory 14 which is a storage cell. The compressed image data recorded on the memory 14 is read via the card I/F circuit 15and is outputted and recorded on IC memory card 16.

[0016]The compressed image data recorded on IC memory card 16 is read via the card I/F circuit 15. The read compressed image data is written in the record section for the pictures of the memory 14. The written-in compressed image data is elongated by real time in the picture expanding part of image data compression / expansion circuit 13. The elongated dynamic image data is supplied to the digital signal processing circuit 12. The luminance signal Y and the color-difference signal Cb which were supplied to the digital signal processing circuit 12and Cr are outputted to LCD(liquid crystal display) 17 as a video signal by which digital decoding was carried out and D/A conversion was carried out to the NTSC signal.

[0017]The operation inside the camera of CPU22 is controlled according to external operation directions of the directions from the final controlling element 21or a remote control (not shown). The display of a camera internal state etc. is displayed on the indicator 20 which comprises LCDLEDELetc.for example.

[0018]An example of the memory area of IC memory card 16 is shown in drawing 2. This IC memory card comprises the boot area M1the memory allocation table region M2the directory information table region M3the file information table M4and the file data area M5. Boot block information is recorded on the boot area M1. To this boot block informationthe version of this IC memory cardThe pointer of a titlea passworda rewriting datea rewriting modelpartition informationa root directoryand a memory allocationthe pointer of size and directory informationthe pointer of size and file informationsizeetc. are recorded.

[0019]Memory-block information is recorded on the memory allocation table region M2. The information on each block that this memory-block information divided the record section of this IC memory card into the block of the plurality of predetermined size is recorded. In for examplethe case so that the file data recorded by the user may be recorded ranging over two or more blocks. When more information which shows whether it is the block to which it was linked from the leading block whether this block is a leading blockinformation on being that this block is usedand this block than the number of erase times set up beforehand are eliminatedthe information which considers the reliability of data and is made impossible [ use ] is recorded.

[0020]Whether it is the number of erase times of this block and the type of this blockfor examplethe block currently used for the boot area. ** is recorded [ whether it is the block with which the file data is recorded / whether it is the block currently used / whether it is the block currently used / whether it is a memory area and / for the directory information table M3and / for the file information table M4and / by the userand ] as information. As mentioned abovewhen there is a block linkedthe address of the following block is recorded as information andas for an end ******** casethe information of an end is recorded

with this block.
[0021]The directory information used as the directory currently used for the directory information table for memory-block information is recorded on the directory information table region M3. In for examplethe case so that the file data recorded by the user may be recorded ranging over two or more directories. The information which shows whether it is the directory to which it was linked from the top directory whether this directory is a headWhen the information on being that this directory is used and the number of the files recorded in this directory are set up beforehandAs the information on the file already recorded is recorded as the element 0 (file number)the element 1 (file number)...the element n (file number) and was mentioned aboveWhen there is a directory linkedthe address of the next directory is recorded as informationand when having ended by this directorythe information of an end is recorded.
[0022]The file information considered as the file currently used for the file information table for memory-block information is recorded on the file information table region M4. For example. [ whether this file is used as a filewhether it is used as a directoryand ] It is recorded whether it is intact andin the case of whether owner IDthe creation date when this file was created or the corrected modification dateand this file are carried out to a read-only file or it is made the file which can be overwrittenand a programit is recorded whether it prevents from performing. When the size of this filethe namethe number of memory block or the number of a directoryand the file are recordedan attributecreator informationetc. on a file are recorded. With this creator informationit is recorded with what kind of applicationfor example this file was created or with what kind of application it is usable in this file.
[0023]The file data currently used for the file data by the user for memory-block information is recorded on the file data M5. As for this file datacompressed image datacompression audio dataetc. are recorded. As an example of these controlling methodsthe absolute address of IC memory card 16 is used.
[0024]In this one examplein image data compression / expansion circuit 13in order to compress dynamic image datacoding/decoding of MPEG 2 are usedand in order to compress voice datacoding/decoding of ATRAC are used in voice data compression / expansion circuit 4. The dynamic image data and voice data which were coded are recorded on the file data M5 of IC memory card 16 mentioned above. An example of the data format is explained. IC memory card 16 is recorded in the format which records / is easy to reproduce with the operation system by the side of a personal computerin order to enable record/reproduction with a personal computer. Generallya DOS format is used. ThereforeDOS file form is applied also in this example.
[0025]HereATRAC is explained briefly. The audio signal of an analog is sampled by a 44.1-kHz sampling frequencythe digitized voice data is cut down by the time window for a maximum of 11.6 msand it decomposes into a frequency component by modification DCT operation. A frequency band is divided into three in order to avoid the preecho which is easy to produce in modification DCT operation. After

changing into a frequency axis by modification DCTdata is operated on a curtailed schedule using human being's aural characteristic. It leaves the ingredient which can be heard since the noise or the small signal near the large music signal frequency component cannot be heard if auditory masking is usedand the ingredient which cannot be heard and which is not required is cut. and compressed data -- 1 sound group -- (-- it bundles to every SG) and is recorded on a recording medium by a cluster unit.

[0026]This ATRAC is adopted by MD (mini disc)and is the method of compression of the voice data standardized. In this MDthe mode 2 of CD-ROM is used and compression audio data is recorded on the magneto-optical disc for sound recording. The data configuration of MD recorded on the magneto-optical disc for sound recording is shown in drawing 3. In this magneto-optical disc for sound recordingas shown in drawing 3 A36 sectors are recorded as solidifying with ** (= one cluster). Three top sectors are made into a link sector region4 sector eye is made into a sub-data sector regionand the 32 remaining sectors are treated as a field for compressed data. As shown in drawing 3 Bone sector consists of 2352 bytesamong those the object for data is 2332 bytes. Compression audio data treats 424 bytes as one unitand calls the unit a sound group. As shown in drawing 3 C11 sound groups are recorded on two sectors. And if 1 sound group's compression is solved and it returns to time base information as shown in drawing 3 Dit will become 512 samples of a right-and-left channel.

[0027]As shown in drawing 4it is a case where gave MPEG 2 to dynamic image data and ATRAC is given to voice dataand a video file and a voice file are considered as another fileand three files of the hour entry management file for taking the synchronization with dynamic image data and voice data become 1 set of data. The video file shown in drawing 4 A comprises a header area and a field of dynamic image data. The record timescreen sizea frame numberrecord bit stream IDpicture achievingetc. of the dynamic image data currently recorded are recorded on the header area of the graphics file as an example. Picture achieving is a ratio of the luminance signal Y and the color-difference signal Cband the sampling frequency of Crfor exampleit is set to 4:2:2 or 4:1:1. Dynamic image data is divided into the GOP units of MPEG in this one example.

[0028]The voice file shown in drawing 4 B comprises a header area and a field of voice data. The record timea sampling frequencya quantifying bit numberrecord bit stream IDetc. of the voice data currently recorded are recorded on the header area of the voice file as an example. Voice data is divided per audio group (AG) who collected the clusters of ATRAC mentioned above.

[0029]And the hour entry management file shown in drawing 4 C comprises a header area and a field of voice data similarly. Data sizedata type IDetc. of a hour entry management file which are memorized are recorded on the header area of a hour entry management file as an example. Hour entry data is time management informationand is recorded by the contents which show below the data address of the voice data which comprised a block of fixed sizefor examplesynchronized with the dynamic image data in GOP units.

[0030]The start address of GOP-N is recorded and AG number (M) which synchronizes with GOP-N is recorded. And the start address of AG-M is recorded and the offset in the AG-M is recorded.
[0031]Concretelyvoice data is divided into an audio group on the basis of GOP of MPEG 2 which compresses into dynamic image data. An example of GOPan audio groupand a hour entry file is shown in drawing 5. As shown in drawing 5 Ain the case of a video fileMPEG 2 is compressed into dynamic image dataand it is recorded by GOP units. Since MPEG 2 of the unfixed length compressing method is usedthe length of the record section of the data corresponding to each GOP differs. Thusdynamic image data is recorded per several frames (GOP units) from which each data size differs.
[0032]As shown in drawing 5 Bin the case of a voice fileATRAC is compressed into voice data and it is considered as a data configuration as shown in drawing 3 B. For examplethe compressed audio group corresponding to 1GOP is unfixed length. Thusvoice data is also recorded by set (audio group) of GOP and a corresponding sampleand this audio group's data sizes also differ.
[0033]And as shown in drawing 5 Cin the case of a hour entry management filethe hour entry for taking the synchronization with GOP of dynamic image data and the audio group of voice data is recorded. As for this hour entry datadata is divided into GOP units. In this examplea hour entry is recorded on a hour entry number (p)andas for a GOP number (n) and the audio group number (m)the start address of the voice data which synchronized with dynamic image data and its dynamic image data for every hour entry number is stored. When 1GOP and the accompanying audio information are 1 audio groupsit becomes the same as that of a GOP number (n) and the audio group number (m)but 1GOP may correspond with two or more audio groups.
[0034]Thusan example which reproduces video and a sound from the IC memory card recorded by the file from which compressed image data and compression audio data differ is explained. Drawing 6 is a block diagram showing an example of the reproduction special-purpose machine of an IC memory card. The hour entry management file by which IC memory card 32 is first reproduced by the card I/F circuit 33 is read. According to the read hour entry management filecompressed image data and compression audio data are read. Compression audio data is supplied to voice data compression / expansion circuit 34and compressed image data is supplied to image data compression / expansion circuit 38.
[0035]In image data compression / expansion circuit 36compressed image data is elongated and the digital signal processing circuit 39 is supplied as dynamic image data. In the digital signal processing circuit 39the supplied dynamic image data is changed into the signal of the analog-ized NTSC systemand is supplied to LCD40. A dynamic image signal is displayed in LCD40. In voice data compression / expansion circuit 34compression audio data is elongated and D/A converter 35 is supplied as voice data. The analog-ized audio signal takes the dynamic image signal and synchronization which are displayed on LCD40 from the loudspeaker 37 via the amplifier 36and is outputted. These control is performed by CPU31.

[0036]An example of control of CPU at the time of this reproduction is explained using the flow chart of drawing 7. In Step S1a hour entry management file is read from an IC memory cardand it is recorded on the memory of CPU22. An input of the position which wishes to reproduce will compute the start address (starting point) of the compressed image data which wishes to reproduce from the hour entry management file recorded on the memory in Step 2. In Step 3the start address (starting point) of the compression audio data reproduced from the hour entry management file currently recorded similarly is computed.
[0037]In step S4compressed image data is read from the start address of the computed compressed image data. The read compressed image data is sent to an elongation processing part.In Step S5compression audio data is read from the start address of the computed compression audio data. The read compression audio data is sent to an elongation processing part. When voice data has offset at this timean offset part shifts voice data and it is elongated. By thisthe synchronization of the dynamic image data and voice data which were reproduced is taken by GOP units.
[0038]In this one examplealthough synchronized on the basis of dynamic image datait may synchronize on the basis of voice data.
[0039]In this one exampleit explained on the assumption that management what is called by a DOS filebut management by the other managing system is realizable similarly.
[0040]
[Effect of the Invention]If it depends on this inventionwhen reproducing the video and sound of arbitrary timethe hour entry file of that time can be readdynamic image data and voice data can be synchronized according to the start address of the dynamic image data and voice data which were recorded on that hour entry fileand it can reproduce.

## DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]
[Drawing 1]It is a block diagram of one example of this invention.
[Drawing 2]It is a schematic diagram of the memory of the IC memory card concerning this invention.
[Drawing 3]It is an approximate line figure used for explanation of ATRAC.
[Drawing 4]It is an approximate line figure used for explanation of the data recorded on the IC memory card concerning this invention.
[Drawing 5]It is an approximate line figure used for explanation of the data recorded on the IC memory card concerning this invention.
[Drawing 6]It is a rough block diagram for explaining an example at the time of reproduction of this invention.
[Drawing 7]It is a flow chart for explaining an example at the time of reproduction of this invention.

[Drawing 8]It is an approximate line figure used for explanation of MPEG.
[Drawing 9]It is an approximate line figure used for explanation of digital VCR.
[Description of Notations]
1 ... A microphone26 ... An amplifier311 ... A/D converter4 ... Voice data compression / expansion circuit5 ... A D/A converter7 ... Loudspeaker8 ... A lensa converging section9 ... A CCD image sensor10 ... CDS circuit12 ... A digital signal processing circuit13 ... Image data compression/expansion circuit14 [ ... LCD18 / ... A driver19 / ... A timing signal generation part20 / ... An indicator21 / ... A final controlling element22 / ... CPU ] ... A memory circuit15 ... A card I/F circuit16 ... An IC memory card17